Numerical Weather Prediction

Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter perturbation strategies on a simple model

Forecasting Research Technical Report No. 453

Neill Bowler

email: nwp_publications@metoffice.com

©Crown Copyright

Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter perturbation strategies on a simple model

Neill E. Bowler

October 12, 2004

Abstract

An experiment has been performed, using a simple chaotic model, to compare different ensemble perturbation strategies. The model used is a 300 variable Lorenz 95 model which displays many of the characteristics of atmospheric numerical weather prediction (NWP) models. 20 member ensembles were generated using five perturbation strategies, error breeding, singular vectors, random perturbations, the Ensemble Kalman Filter (EnKF) and the Ensemble Transform Kalman Filter (ETKF). The EnKF and random perturbations are clearly superior to the other methods; they are the only methods to show growth of the ensemble spread at broadly the correct rate for the entire forecast period and have a rank histogram which is very close to flat. The ETKF ensemble is seen to be superior to the singular vectors ensemble, which in turn has better performance than the error breeding ensemble.

Central to the success of the EnKF is the localisation of the background error covariance which removes spurious long-range correlations estimated by the ensemble. Computationally efficient versions of the EnKF (such as the ETKF) cannot accommodate covariance localisation and their performance is seen to suffer. Combining random perturbations with either the ETKF or error breeding produces an ensemble which is superior to any of the these three ensemble generation techniques considered alone. Therefore, "stochastic physics" schemes may considerably improve the performance of the initial conditions perturbations for methods which rely on a filtering approach. Applying the ETKF to a series of local domains has been tested, which allows covariance localisation whilst remaining computational efficient, and this has been found to be nearly as effective as the EnKF with covariance localisation.

The limitations of testing on this simple model without model error are evident since random, unbalanced perturbations can produce a near-perfect ensemble. Hence these tests can only provide an indication of possible performance with NWP models.

1 Introduction and motivation

The Met Office intends to develop a short-range ($<$ 3 days) ensemble prediction system based on a limited area model (LAMEPS). This study aims to provide guidance on the development of initial conditions perturbations for that model using a simple model to test various strategies.

Ensembles provide a pragmatic approach to sampling the posterior pdf of a numerical weather prediction (NWP) model forecast. The uncertainties in the forecast are often partitioned into those arising from uncertainties in the initial conditions and uncertainties in the model formulation, and here the focus is on the former of these two.

Over the years a number of methods for the generation of perturbations to the initial conditions have been developed, such as error breeding [21, 22], singular vectors [4, 18], the ensemble Kalman filter (EnKF) [6], perturbed observations [11] and random perturbations [12, 5]. From these methods it is necessary to choose one which will provide adequate performance for the proposed LAMEPS system. Recent studies [25, 3] have aimed to compare global EPS's from different centres, but it has proved difficult to disentangle the quality of the NWP model from the quality of the ensemble generation system. Nonetheless, these studies suggest that error breeding may be superior to singular vectors for short lead times. Studies on the ensemble transform Kalman filter (ETKF) have indicated that this is superior to error breeding [23]. Hence, it would seem that the ETKF is the perturbation strategy to use for the LAMEPS.

Hamill *et al.* [9] compared error breeding, singular vectors and perturbed observations on a quasi-geostrophic channel model. Perhaps unsurprisingly the perturbed observation approach was close to ideal, although this approach is known to be computationally very costly. The present study is similar to that of Hamill *et al.* [9], but the focus will be on the EnKF and ensemble transform Kalman filter (ETKF) and whether these are superior to the more established methods of error breeding and singular vectors. This test is based on the Lorenz 95 model [15, 16]. Care has been taken to implement a scheme where all the perturbation strategies are treated equally. The model used is considerably simpler than NWP models and therefore any results cannot be directly applied to NWP ensembles. However, results from these tests should provide further guidance on the perturbation strategy to be used at the Met Office.

2 Lorenz 95 Model

The Lorenz 95 (L95) model [15, 16] is defined by the following system of differential equations

$$\frac{dX_i}{dt} = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F \tag{1}$$

where the variables X_i, $i = 1, 2, ..., N$, are defined on a cyclic chain such that $X_{-1} = X_{N-1}$, $X_0 = X_N$ and $X_1 = X_{N+1}$. Different behaviour is seen depending on the magnitude of the forcing, F. For small values of F the system converges

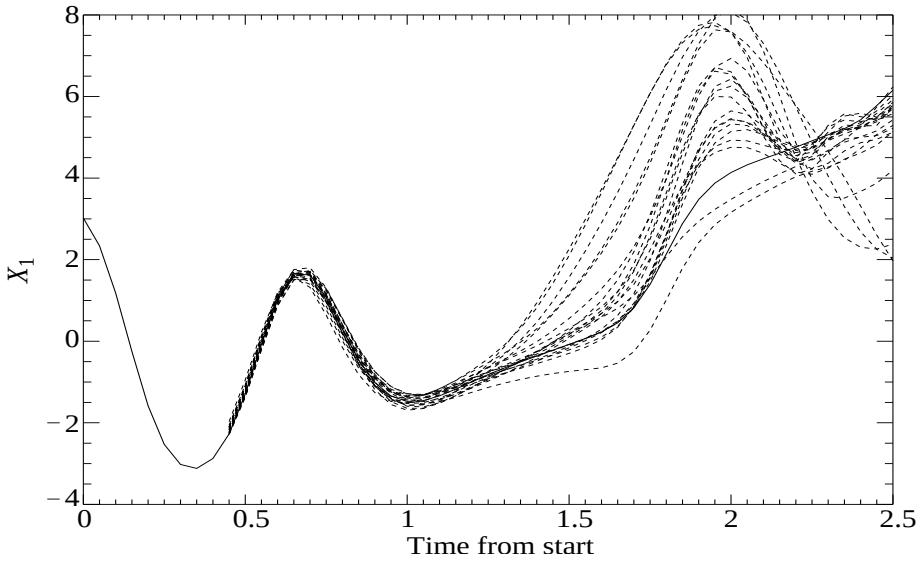

Figure 1: Example trajectory of the first variable X_1 for the truth run (solid line) and the 20 ensemble members calculated using the EnKF (dashed lines). At all points the truth lies within the spread of the ensemble, and the ensemble spread increases with increasing lead time.

to a fixed point ($X_i \to 0$ for all i when $F = 0$) whereas the system is chaotic for large values of F. In all the experiments here $F = 8$ is used, since this is well into the chaotic regime. The Runge-Kutta 4^{th} order method was used to perform the time stepping, for intervals of $\delta t = 0.05$. The system size is chosen to be $N = 300$, since this allows an ensemble of size $K = 20$, whilst still maintaining the ensemble size much less than the system size. Figure 1 shows an example ensemble forecast using the ensemble Kalman filter. The solid line shows the true trajectory which the ensemble is attempting to capture, and this lies within the spread of the ensemble (dashed lines) for all lead times.

3 Analysis method

The Lorenz model was run with a time step of $\delta t = 0.05$ for 2000 time steps, the first 1000 being used to allow the system to "spin-up". Every time step observations were assimilated using a robust assimilation method which is independent of the ensemble perturbations. After each occasion of observations being assimilated an ensemble forecast with 20 members was run for 50 time steps. Hence a set of 1000 ensemble forecasts has been produced, with these forecasts being verified against a "truth run".

The observations were a perturbed version of the state of the "truth run" with an error standard deviation of 0.1. A set of observations contains the values of all

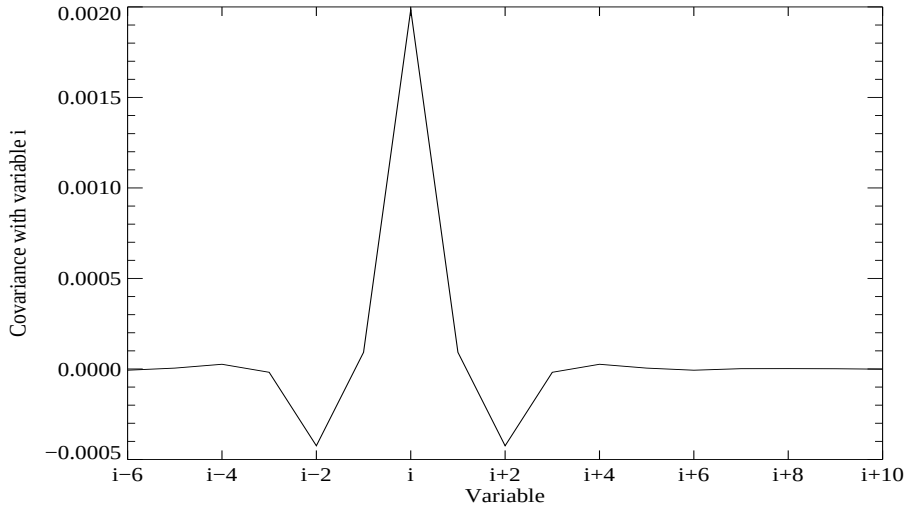

Figure 2: Example row from the constant background error covariance matrix used in the 3D-Var analysis scheme. The matrix is symmetric and essentially zero for elements more than two variables away from the diagonal.

300 state variables. The analysis update is performed using the 3D-Var method, but an exact minimisation is performed using the serial updating approach (see EnKF section 4.3). A control forecast is run, and this is updated by computing

$$\mathbf{X}^{ca} = \mathbf{X}^{cf} + \mathbf{P}^b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R}\right)^{-1}\left(\mathbf{Y} - \mathbf{H}\mathbf{X}^{cf}\right) \tag{2}$$

where $\mathbf{X}^{ca}$ is the control analysis, $\mathbf{X}^{cf}$ is the control forecast, $\mathbf{H}$ is the observation operator (the identity here), $\mathbf{P}^b$ is the background error covariance, $\mathbf{R}$ is the observation error covariance (diagonal matrix with elements equal to 0.01), and $\mathbf{Y}$ are the observations. $\mathbf{P}^b$ is assumed constant, as is normal in operational NWP forecasts, and has been estimated from a long run of an EnKF ensemble. A row from this matrix is shown in Fig. 2 which is symmetric and essentially quin-diagonal[1]. $\mathbf{P}^b$ was also estimated from a long run of a singular vectors ensemble, but this resulted in an inferior analysis.

It is possible to improve this data assimilation by allowing a time-varying $\mathbf{P}^b$ (see section 5.1 below). However, a constant $\mathbf{P}^b$ is used as this is what is used for operational NWP and it provides an analysis which is independent of the ensemble quality.

[1] A quin-diagonal matrix contains non-zero entries only in the five elements that are nearest the diagonal in each row

4 Perturbation Strategies

4.1 Error Breeding

To generate an ensemble using error breeding [21, 22], random perturbations are allowed to grow through model integrations. At each assimilation time the perturbations are scaled down by an appropriate factor. This factor is chosen so that the spread of the perturbations to the analysis is a given constant, the size of this constant being tuned so that the forecast ensemble spread at the assimilation time is equal, on average, to the RMS error of the ensemble mean at that time. For each assimilation time, the first 10 perturbations are re-scaled by the appropriate factor. The second 10 perturbations are chosen to be the negative of the first 10, thus ensuring the perturbations are centred around the analysis.

4.2 Singular Vectors

The singular vector (SV) perturbations are found by calculating the fastest growing modes of a linearised version of the dynamics. The linearised version of the Lorenz 95 model is

$$\frac{d\,\delta X_i}{dt} = X_{i-1}(\delta X_{i+1} - \delta X_{i-2}) + \delta X_{i-1}(X_{i+1} - X_{i-2}) - \delta X_i \qquad (3)$$

where terms of order δX^2 and above have been neglected. From this, the matrix transformation for the linearised dynamics may be obtained

$$\lim_{\delta t \to 0} \frac{\delta \mathbf{X}(t+\delta t)}{\delta t} = \mathbf{L}\,\delta \mathbf{X}(t). \qquad (4)$$

$\mathbf{L}$ describes the instantaneous rate of change of a perturbation, and must be combined with $\mathbf{L}$ from different times to account for the Runge-Kutta 4th order time stepping that is used to update the model. This propagator matrix is found for each time step up to a lead time of 0.4 time units (believed equivalent to 2 days of an atmospheric model). These are multiplied together to give a final matrix $\mathbf{M}$ which describes the evolution of the perturbations over the next 0.4 time units. In accordance with the ECMWF system [18] the right singular vectors of $\mathbf{M}^T\mathbf{M}$ are found. Unlike SV's for NWP the linearisation of the model equations is exact, but in all other respects the SV's should be similar to those calculated operationally at ECMWF.

10 of these SV's are chosen, using the overlap function of Molteni *et al.* [18] which aims to exclude SV's which are very similar. Since the SV's tend to be highly localised, the first 10 SV's are typically selected. The SV's are ordered by their singular values which means that a Euclidean norm is used here. Since all the variables are equivalent this is the appropriate norm to choose. The SV's are mixed using a rotation matrix, $\mathbf{S}$, so that the perturbations become

$$\mathbf{X}^a = \mathbf{VS} \qquad (5)$$

where the columns of $\mathbf{V}$ contain the SV's and the columns of $\mathbf{X}^a$ contain the perturbations to be added to the ensemble mean analysis. $\mathbf{S}$ is chosen to minimise the cost function

$$C_f = \sum_{j=1}^{K} \left(\sum_{i=1}^{N} \left(X_{j,i}^a \right)^8 \right)^{1/4} \tag{6}$$

where $X_{j,i}^a$ is the resultant perturbation of variable i for member j. This minimisation is performed to produce perturbations that are similar to the SV's, but have smaller local maxima and a more uniform spatial distribution. C_f is minimised by examining all the pairs of SV's and minimising C_f for each pair in turn. The resultant perturbations are scaled so that the ensemble spread at the assimilation time is equal to the average RMS error of the ensemble mean forecast. The second 10 perturbations are set equal to the negative of the first 10 perturbations. This approach is chosen to closely mimic the approach taken by ECMWF [18].

4.3 Ensemble Kalman Filter (EnKF)

The Kalman filter [2, 14] is a data assimilation scheme which provides an optimal estimate of the true state of the system, given certain conditions. The filter works by maintaining an estimate of the true state, which is updated by a forecast of the state from the previous time and by observations. Errors in the observations, the state at the previous time and the model for propagating this state forward in time are accounted for (although here a perfect model framework is being used). The filter provides an optimal data assimilation system if the model and observation operator are linear, and all the errors are Gaussian. These conditions are not met by NWP models of the atmosphere. For NWP models the ensemble Kalman filter is a useful approximation of the Kalman filter, which allows the data assimilation to be computed on the basis of a limited size ensemble. It still assumes that all the errors are Gaussian (which will not be the case when non-linear models are used) but behaves well even when they are not.

As input the EnKF requires an ensemble of forecasts of the current state from which the background error covariance may be estimated

$$\hat{\mathbf{P}}^b = \frac{\mathbf{X}^f \mathbf{X}^{fT}}{K-1} \tag{7}$$

where K is the size of the ensemble and the columns of $\mathbf{X}^f$ correspond to the difference between an ensemble member and the ensemble mean forecast. In common with others [13] a localisation function is applied to this covariance matrix

$$P_{i,j}^{f,l} = P_{i,j}^f C(|i-j|, D) \tag{8}$$

where C is the popular 5^{th} order function of Gaspari & Cohn [8] and the choice of half-width of the function, D, is discussed in section 5.1, below. This excludes long range correlations in the background error, which are likely to be spurious. The ensemble mean analysis is calculated using

$$\bar{\mathbf{X}}^a = \bar{\mathbf{X}}^f + \mathbf{G}\left(\mathbf{Y} - \mathbf{H}\left(\bar{\mathbf{X}}^f\right)\right) \tag{9}$$

where the symbols take the meanings given in eq. 2, and $\mathbf{G}$ is known as the Kalman gain, given by

$$\mathbf{G} = \hat{\mathbf{P}}^{b,l}\mathbf{H}^T \left(\mathbf{H}\hat{\mathbf{P}}^{b,l}\mathbf{H}^T + \mathbf{R}\right)^{-1}. \tag{10}$$

The inversion of the term on the right of eq. 10 can be very expensive for NWP models which use large numbers of observations.

Equation 9 is correct for performing the update of the ensemble mean, or analysis, but to derive the correct distributions of analysis perturbations it would normally be necessary to perturb the observations, $\mathbf{Y}$. However, since $\mathbf{R}$ is diagonal it is entirely appropriate to process the observations serially [26]. In this case $\mathbf{H}\hat{\mathbf{P}}^{b,l}\mathbf{H}^T$ and $\mathbf{R}$ reduce to scalars. To update the perturbations without perturbing the observations $\mathbf{G}$ is multiplied by

$$\alpha = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\hat{\mathbf{P}}^{b,l}\mathbf{H}^T + \mathbf{R}}}\right)^{-1}. \tag{11}$$

This algorithm performs better than the standard EnKF with perturbed observations [7], and is much faster for the simple model used here.

A fixed inflation factor is applied to the analysis perturbations and along with covariance localisation helps to ensure that the perturbations have the correct spread at the next time step. The choice of the size of these factors is discussed in section 5.1, below. The variable inflation factor of Wang & Bishop [23] would have been used. However, the variable inflation factor did not behave stably, and made the results much worse. This is due to the algorithm introducing a feedback into the perturbation size, creating undesirable oscillations.

4.4 Ensemble Transform Kalman Filter (ETKF)

The ensemble transform Kalman filter (ETKF) [1] is equivalent to the EnKF and allows for rapid calculation of the analysis perturbations. The ETKF is approximately 10^5 times faster than the serial EnKF for current NWP models. The efficiency of the computation relies on the analysis perturbations being in the subspace of the forecast perturbations. Hence, covariance localisation is not possible within the ETKF framework. The scheme centres around calculating a transform matrix, $\mathbf{T}$, which transforms the forecast perturbations according to

$$\mathbf{X}^a = \mathbf{X}^f\mathbf{T}. \tag{12}$$

The analysis perturbations for a given variable are therefore a linear combination of the forecast perturbations for that variable. Thus, the ETKF may be seen as being closely related to the error breeding method.

The transformation matrix is calculated as [23]

$$\mathbf{T} = \mathbf{C}(\Gamma + \mathbf{I})^{-1/2}\mathbf{C}^T \tag{13}$$

where $\mathbf{C}$ and Γ are derived from the eigenvectors and eigenvalues of the matrix

$$\mathbf{E} = \frac{\left(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{X}^f\right)^T\left(\mathbf{R}^{-1/2}\mathbf{H}\mathbf{X}^f\right)}{K-1}. \tag{14}$$

Since $\mathbf{X}^f$ contains the ensemble perturbations (and therefore must sum to zero) one of the eigenvalues of $\mathbf{E}$ will be zero. Γ is a $(K-1) \times (K-1)$ diagonal matrix containing all the eigenvalues of $\mathbf{E}$ except this zero eigenvalue, and $\mathbf{C}$ contains the corresponding eigenvectors, save the one relating to the zero eigenvalue. The perturbations thus generated will be orthogonal to each other [24] and have the appropriate variance, as dictated by the Kalman filter formulation.

4.5 Random perturbations

The final perturbation scheme considered here is the simplest - adding white noise to the analysis. The perturbation at one point is independent of the perturbation at all other points, and conditioned according to

$$\sum_{j=1}^{K} X^a_{j,i} = 0 \qquad \text{for all } i \tag{15}$$

and

$$\sum_{j=1}^{K} \left(X^a_{j,i}\right)^2 = \sigma_n^2 \qquad \text{for all } i \tag{16}$$

where σ_n^2 is chosen so that the spread of the ensemble is on average equal to the RMS error of the ensemble mean forecast at the assimilation time. The distribution of the noise is chosen to fit the distribution of the measured analysis error. This distribution is given is figure 3 and is the convolution of a Gaussian and a symmetric exponential distribution.

5 Results

5.1 Performance of the analysis

The 3D-Var analysis scheme (equation 2) is used when comparing the perturbation schemes. The average error of this analysis over a long forecast cycle is 0.0400. However, if the EnKF scheme is used to update the ensemble mean (as well as updating the perturbations) the analysis error can be as low as 0.0171. This value was achieved with a covariance localisation half-width of 24 variables, and the perturbations were inflated by 1%. If the background error covariance used in the 3D-Var scheme is derived from the EnKF ensemble, then the mean analysis error can be as low as 0.0174, with a covariance localisation half-width of 20 variables. The difference between using the EnKF analysis and 3D-Var with background errors determined from the ensemble is probably not significant. This provides evidence that the current operational analysis schemes could be dramatically improved with the use of ensembles. However, it is likely to prove difficult to obtain adequate estimates of the background error covariance for NWP models which have a considerably higher dimension than this simple model.

Further evidence for the utility of flow-dependent background error covariance estimates comes from the use of small ensembles. An ensemble Kalman filter with

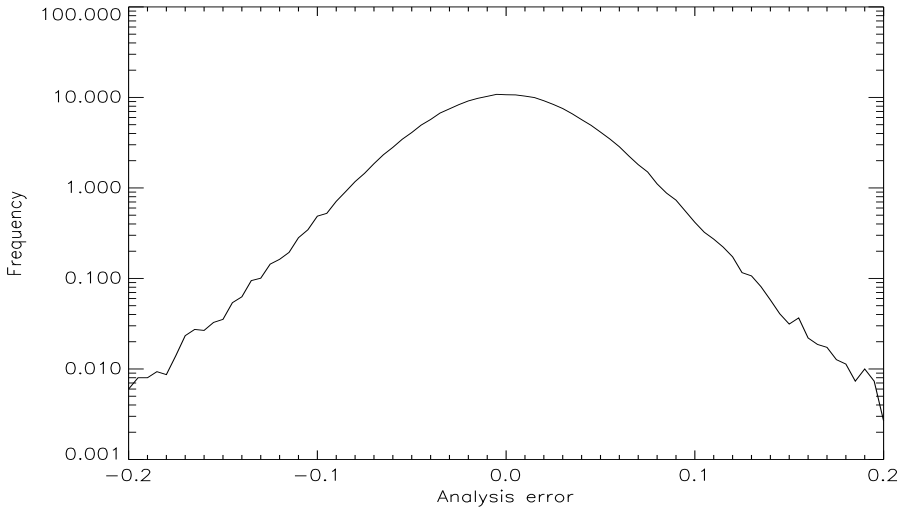

Figure 3: Frequency distribution of the observed analysis errors for the 3D-Var system. This distribution is approximately the convolution of a Gaussian with a symmetric exponential distribution, indicating that large errors in the analysis are more common than in a Gaussian distribution.

3 ensemble members can still provide a superior analysis to 3D-Var with constant background error covariances. Using a covariance localisation half-width of 2.5 variables and inflation factor of 1.1 an average analysis error of 0.0372 was found, slightly superior to the 3D-Var estimates.

When running the EnKF with the 3D-Var analysis scheme, it is chosen to keep the covariance localisation the same as for the optimal performance of the EnKF analysis scheme, and simply increase the inflation of the perturbations to maintain the correct spread in the ensemble. Therefore the EnKF was run with a localisation half-width of 24 variables and an inflation factor of 1.152. Since the ETKF cannot use covariance localisation a larger inflation factor of 1.883 is required to maintain the correct spread of the ensemble.

It is not envisaged that the ETKF perturbations will be used to specify the background error covariance estimates used by the analysis scheme. Even when $\mathbf{P}^b$ thus attained is filtered using covariance localisation the estimates are still deficient and it has proved impossible to form a working 3D-Var & ETKF analysis scheme. Therefore, it would appear that without some kind of intervention (e.g. hybridisation [10], the addition of noise, or some localisation, see below) the ETKF does not provide a good basis for improving analysis schemes.

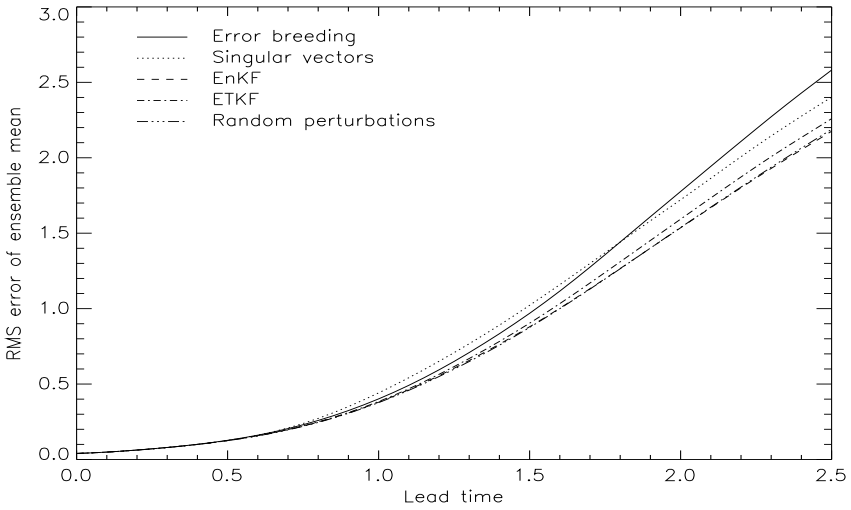

Figure 4: RMS error of the ensemble mean. The EnKF and random perturbations have the lowest errors, performing slightly better than the ETKF. The singular vectors ensemble performs worst initially, due to the ensemble being over-spread (see figure 5) and the error breeding ensemble performs worst at long lead times.

5.2 Performance of the ensemble perturbations

5.2.1 Skill of the ensemble mean

The greater the ability of an ensemble forecast to account for the likely errors in the forecast, the greater will be the skill of the mean of that ensemble forecast. Therefore the skill of the ensemble mean can be taken as a (crude) measure of the quality of the ensemble. Figure 4 shows the RMS error of the ensemble mean forecast for the five perturbation strategies. The skill of the ensemble mean forecast for the EnKF and random perturbations is superior to the skill of all the other methods. The skill of the ensemble mean for the error breeding is superior to the ensemble mean of the SV ensemble at lead times around 1.7 time units. This is probably due to the SV ensemble being over-spread at these times (see below).

5.2.2 Spread - Error of the ensemble

Ideally, the spread of the ensemble will be equal to the RMS error of the ensemble mean at all lead times. Figure 5 shows the RMS spread of the ensemble, divided by the RMS error of the ensemble mean. All of the ensembles have been tuned so that this ratio is equal to unity at the assimilation time (0.05 time units). Both the EnKF and ETKF show robust performance with the ensemble spread close to one for all lead times. There is little difference between these filters for up to lead time

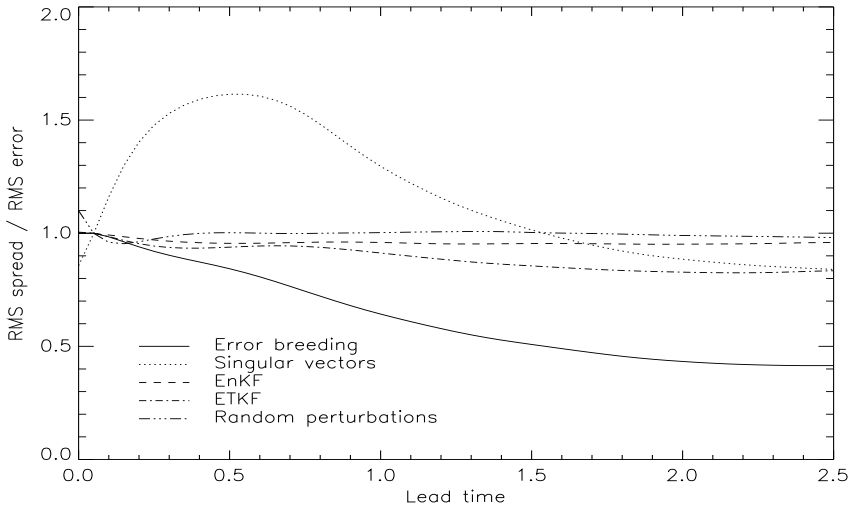

Figure 5: Ratio of the RMS error of the ensemble mean to the spread of the ensemble. All of the schemes are tuned to have the correct spread at time step 3. For the ensemble Kalman filters this ratio stays close to 1 for all lead times. However, the ensemble spread for the error breeding and singular vectors methods does not grow at the correct rate. The full EnKF appears to be slightly superior to the ETKF.

of 0.8 time units, with the EnKF slightly superior. After this time the performance of the ETKF ensemble deteriorates further. Initially the spread of the random perturbations ensemble decreases due to the perturbations being unbalanced, but then levels off to a value very close to one.

Error breeding is clearly the worst performer amongst the schemes, showing growth that is much too slow. The poor performance is attributable to the perturbations being very similar, as is seen later (sec. 5.2.4). The initial growth of spread of the SV ensemble is rapid. At around the optimisation time (0.4 time units) the growth rate plateaus, after which the growth rate is too slow.

5.2.3 Rank Histogram

If each ensemble member is an equally probably sample from the true pdf of the forecast, then the truth should have the same statistical properties as the ensemble members. For each variable one can order the ensemble members, and find where the truth lies within this order. If the ensemble forecast samples the true forecast pdf, then the truth is equally likely to lie in any of the $K+1$ bins.

The error breeding and SV's schemes do not aim to provide a random sampling of the forecast pdf, rather they attempt to identify which errors are likely to grow rapidly. Computation of a pdf from the ensemble forecast is a non-trivial work.

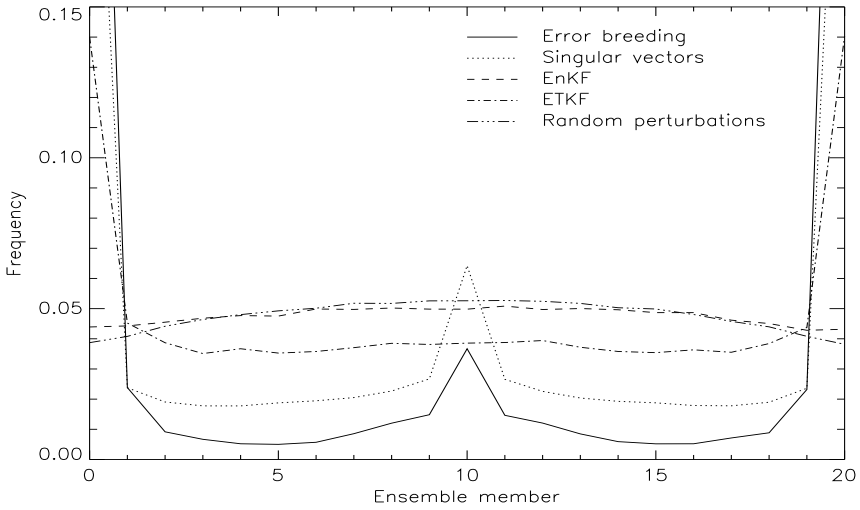

Figure 6: Rank histogram for the four perturbation strategies, evaluated for forecasts at the assimilation time. Error breeding is clearly inferior to the other methods, and the full EnKF appears to have the best performance. The peak in the centre of the diagram for singular vectors and error breeding is due to the use of $+/-$ pairs in centring the perturbations.

Nonetheless, the rank histogram can provide useful information on the quality of the ensemble schemes. In particular, if the number of times the truth lies outside the range of the ensemble is larger than would be expected, then this is a clear sign of problems with the ensemble.

The rank histogram for forecasts at the assimilation time (fig. 6) indicates that the EnKF is superior to the other methods with its rank histogram being closer to flat than the rank histograms for the other methods. The rank histogram for the random perturbations has too few outliers, indicating that the perturbations are less similar to each other than a random draw from the pdf of the forecast. This is because the ensemble mean of the perturbations is set to zero at every point (see eq. 15) which introduces a small anti-correlation between the perturbations. There is a peak in the centre of the rank histogram for the SV and error breeding perturbations which is a consequence of choosing $+/-$ pairs to centre the ensemble. The ETKF performance is superior to error breeding and SV's - even though all the ensembles have the same spread at this time.

A crucial aspect of rank histograms is the frequency with which the truth lies outside the spread of the ensemble. This frequency is a single-variable estimate of the skill of an ensemble, and the frequency of outliers, relative to the ideal value of $\frac{2}{K+1}$ is plotted in figure 7. Error breeding has a large number of outliers for all lead times. The SV's ensemble has a quick reduction in the number of outliers, due to the ensemble becoming over-spread. This doubly indicates the failings of

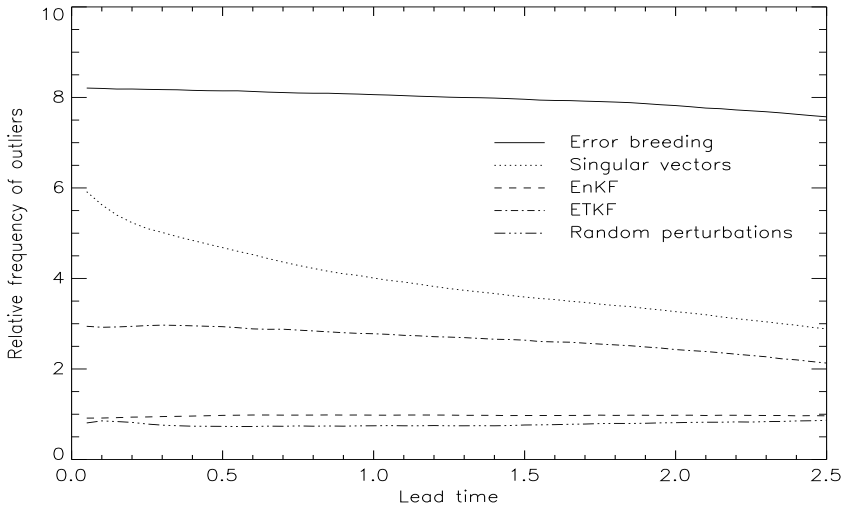

Figure 7: The frequency of the excessive number of outliers for varying lead times. The EnKF ensemble is close to perfect, and for the other methods the truth too often lies outside the range of the ensemble. As the spread of the SV's ensemble increases very rapidly initially, it appears to improve quickly with increasing lead time.

the SV's ensemble, since over-spread ensembles would usually be associated with too few outliers. The ETKF varies between having two to three times the ideal number of outliers, which is always less than for error breeding or SV's. The EnKF has slightly too few outliers, but is closer than all the other methods to the ideal frequency for all lead times. The random perturbations ensemble has even fewer outliers, for the reason discussed above. All of the ensemble forecasts have closer to the ideal number of outliers at longer lead times indicating that performance at short lead times is most revealing. This is a general feature of ensemble forecasts, as noted by Wei & Toth [25].

5.2.4 Average eigenvalues

The average eigenvalues of the normalised forecast perturbations (fig. 8) can provide useful information on the independence of the perturbations from one another [23]. The sum of these eigenvalues is approximately equal for all the perturbation schemes, since the spread of the ensemble at the assimilation time is approximately the same. For error breeding there is a great deal of similarity between the ensemble perturbations, indicated by the first eigenvalue being much larger than all the others. The singular vector perturbations show a much more even distribution of the eigenvalues, but the second 10 perturbations are very similar to the first 10, as might be expected. The EnKF, ETKF and random perturbations generate a broad

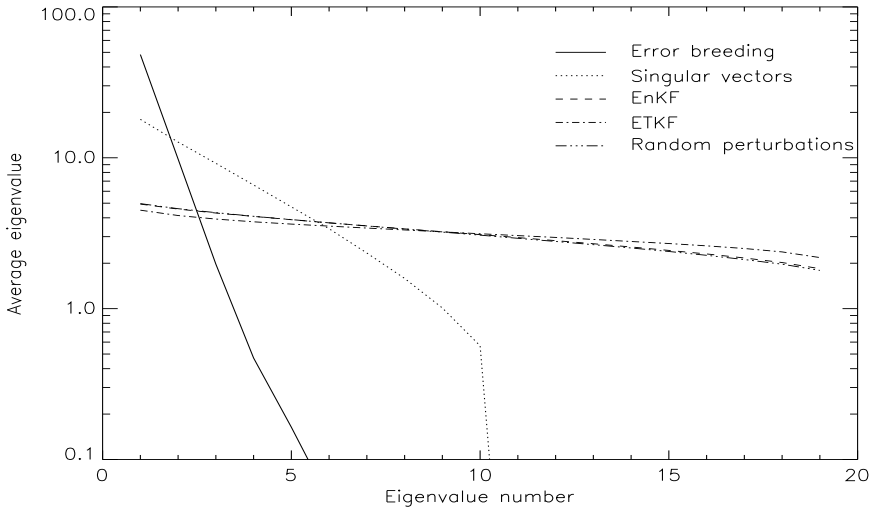

Figure 8: Sum of the eigenvalues of the forecast perturbations in normalised observation space (see [23]). This indicates the extent to which the forecast perturbations are independent of each other. The SV's method uses $+/-$ pairs for centring the ensemble, and it can be seen that one half of these pairs is very similar to the other half. Both ensemble Kalman filters show good independence of the perturbations.

spectrum of independent perturbations, with increased independence for the ETKF perturbations, presumably due to the orthogonalisation that has been applied to this set.

6 Discussion

6.1 Performance of singular vectors, error breeding and random perturbations

In this study the two most established methods of initial condition perturbations appear to be the worst performing methods. For error breeding this is due to the perturbations being very similar (as is seen in figure 8). It is less clear why the SV's ensemble performs poorly.

The SV's identify those structures which will grow most rapidly during the forecast. However, there is no guarantee that the analysis error projects significantly onto these vectors. The rank histogram suggests (figure 6) that either the singular vectors are often not similar to the analysis error, or that the choice of only the first 10 singular vectors is insufficient.

Apart from the 'n' shaped rank histogram, the random perturbations ensemble

appears to perform extremely well. Since the average background error covariance has nearly all its weight at the origin, the errors in the analysis are typically uncorrelated, and hence white noise would be expected to closely reproduce the errors observed in the analysis. This, and the insensitivity of the L95 model to imbalance, mean that the random perturbations ensemble performs well. Since NWP models do not share the above characteristics, a random perturbations ensemble would not be expected to perform so well for NWP models.

6.2 Adding noise to the perturbations

If it is chosen to run the EnKF without localisation then one may take advantage of the considerable computational speed increase of the ETKF. However, it has been shown above that covariance localisation brings considerable improvements to the ensemble, especially when the background error covariance estimate is used in an analysis scheme.

The effect of covariance localisation is to ensure that the analysis update for distant points in the system are performed independently. Therefore, it is expected that the average background error covariance, $\mathbf{P}^b$, between distant points will be zero (as is the case, see figure 2). However, due to random sampling using a limited size ensemble the root mean square (RMS) of the background error covariance between distant points should be $\frac{1}{\sqrt{K}}$ times the average value of the background error variance. This is seen for the EnKF with localisation (figure 9), which also shows a noticeable suppression of the variability in $\mathbf{P}^b$ within the localisation area. It may be possible to improve the localisation by fitting a correlation function to $\mathbf{P}^b$, rather than multiplying by another function. However, this would require knowing the appropriate function to which to fit.

The ETKF (or equivalently EnKF without localisation) also provides a good estimate of the average background error covariance between distant points (indistinguishable from figure 2). However, the variability in $\mathbf{P}^b$, as measured by the RMS, is much reduced for distant points and much enhanced for nearby points, see figure 9. Therefore, the spurious long-range correlations introduced by the small sample size are squashed by the EnKF processing, reducing the variability in $\mathbf{P}^b$ at long range. The effect of localisation is to restore the variability in the background error covariance estimated from the ensemble. This squashing of the variability may be the effect responsible for the suppression of the variability for the EnKF with localisation within the localisation area.

To improve the quality of the ETKF perturbations without the cost of localisation, it is considered here to combine the ETKF with random perturbations (RP's). Since RP's are uncorrelated they may increase the variability of $\mathbf{P}^b$ at long range for the ETKF. Figure 10 shows $\mathbf{P}^b$ for the ETKF combined with random perturbations and how this can closely resemble the shape one would expect.

Since the performance of the random perturbations ensemble is near-perfect it is difficult to quantify the improvement to the individual schemes by combining them with RP's. The quality of the analysis that is produced when using the ensemble estimate of $\mathbf{P}^b$ in the 3D-Var scheme is used to measure the quality of the ensemble

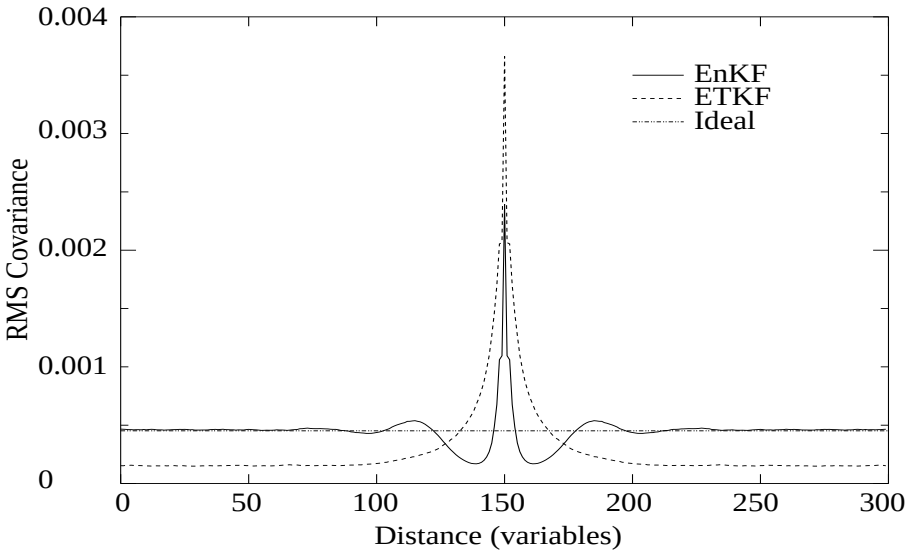

Figure 9: The RMS of the background error covariance between a given point and variable number 150, $\sqrt{\overline{(\mathbf{P}^b)^2}}$, estimated from the EnKF (solid line) and ETKF (dashed line) ensembles. The dotted line shows the value that would be expected for the EnKF at long distances, based on the updates of distant points being performed independently. The value of this quantity for distant points from the ETKF ensemble is suppressed due to spurious long-range correlations in the ensemble.

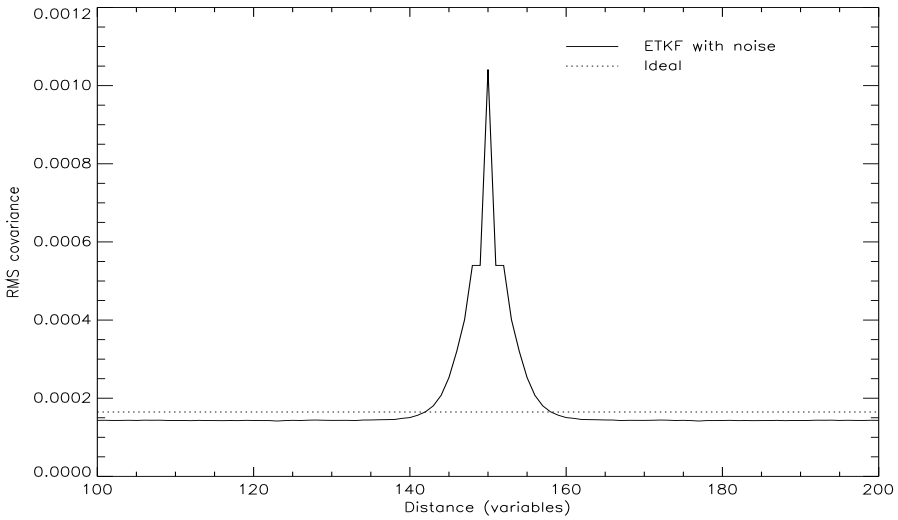

Figure 10: The RMS background error covariance with variable 150 for the ETKF with added noise, parameters tuned as in table 1. Also plotted is the ideal value this would take if the analysis update for distant points were independent. The value seen for the ETKF is still slightly suppressed.

forecasts. Earlier it was stated that using $\mathbf{P}^b$ estimated from the EnKF gave an improvement to the analysis of greater than 50%, which is used as the benchmark for the quality of the combined ETKF-random perturbations ensemble.

Although the ETKF is incompatible with covariance localisation, the $\mathbf{P}^b$ estimated by the ETKF may be localised before it is used as input to the analysis scheme. The best results for each of the perturbation schemes are given in table 1. This indicates that the EnKF with covariance localisation is the best perturbation strategy, followed by the combined ETKF-random perturbations scheme. For this scheme, the RP's contribute around 7% of the variance of the total perturbations. The performance of error breeding is much improved by combining with RP's. This may explain why error breeding has not been seen to be clearly inferior to SV's for NWP models, since deterministic parameterisations with their numerous switches may provide some of the noise required to keep the bred modes distinct from each other. Combining SV's with RP's did not provide an improved analysis compared with the RP's alone. This indicates that the singular vector perturbations, even when combined in the complex way described in section 4.2 do not provide an estimate of the analysis error.

It is somewhat counter-intuitive that the performance of an ensemble forecasting system can be improved by adding noise, especially within a perfect model environment. The success of the RP scheme and combining this with other schemes is likely to be limited to simple models. For NWP, model error has to be considered and one approach to this is a "stochastic physics" scheme (where the forecast model itself

Method	Noise σ_n	Localisation D	Mean spread	Analysis error
Static P^b	0	N/A	N/A	0.0400
EnKF	0	24	0.0173	0.0174
ETKF + RP	0.0064	7.5	0.0241	0.0249
Breeding + RP	0.0120	5	0.0277	0.0296
RP	0.0453	3.5	0.0400	0.0369

Table 1: The best analysis which can be found for the various perturbation schemes when combined with the random perturbations (RP's). The EnKF with localisation is clearly superior to the other methods. Combining singular vector perturbations with RP's produced an analysis which was inferior to using RP's alone.

is allowed to evolve with time). Such schemes may be a very convenient way to add balanced noise to the perturbations, improving the performance of the ETKF.

6.3 Localisation and the ETKF

Another approach to improve the ETKF perturbations is to apply the ETKF over a series of local areas, similar to what has been proposed by Ott *et al.* [19]. They calculate the analysis for a given grid point as the solution of the EnKF over an area centred on that grid point. Since only observations and model forecasts local to that point are considered in updating the ensemble at the point in question, this is an effective form of covariance localisation. Also, since the area over which the EnKF is solved only changes one grid-point at a time, the solution produced by this algorithm will vary smoothly over the model domain. However, this scheme is still potentially rather expensive, since it requires that the EnKF is solved for every grid point.

With the ETKF, a rather helpful enhancement to the local EnKF scheme presents itself. The solution of the ETKF gives a transform matrix, in which the analysis perturbations for some variable at a point are written as a linear combination of the forecast perturbations for that variable at the same point. By solving the ETKF over local areas one may produce the same solution as for the local EnKF scheme. However, it is possible to solve the ETKF at a sparse set of points, and interpolate the transform matrix between these points. The difficulty is that the transform matrix calculated at one point needs to be consistent with the transform matrix calculated for all adjacent points. The determination of the transform matrix by the spherical simplex method (see above and [24]) may give the required consistency, since this necessarily requires that all the diagonal elements of the matrix are positive. Provided the areas over which the ETKF is solved are sufficiently large and the points are not too sparse, then the interpolation should not cause serious problems.

To demonstrate the effect of the interpolation of the transform matrix, the local ETKF scheme is run using an area size of 41 variables (very similar to a localisation

Number of points	Localisation D	Mean spread	Analysis error
300	20	0.0182	0.0179
100	20	0.0182	0.0178
60	20	0.0182	0.0178
20	20	0.0181	0.0177
10	20	0.0181	0.0179

Table 2: Performance of the 3D-Var analysis scheme when the local ETKF scheme is used to specify the background error covariance. For all configurations local areas of 41 variables were used, and covariance localisation applied to the $\mathbf{P}^b$ calculated from the ensemble. The interpolation does not appear to adversely affect the system, and the differences in performance are not believed to be significant.

half-width D of 20 variables) for various numbers of points. Where the scheme is run for 300 points this corresponds to the original local EnKF scheme, and where it is run for 150 points the transform matrix is calculated at every other point. The background error covariance matrix from this ensemble is calculated, localised in the standard manner, and used by the 3D-Var scheme. Table 2 gives the performance of the analysis when the local ETKF approach is used. There is little difference between the approaches for all the number of points used. If the local areas of adjacent points do not overlap, then the scheme is unstable, but it appears to work well in all other circumstances. Furthermore, since the quality of the analysis does not deteriorate with increased interpolation these results indicate that the transform matrix obtained in one region is very similar to that obtained in the adjacent regions. Given as areas for the local EnKF with an NWP model are given at around 1000km across for an ensemble of size 40 [20], the local ETKF could safely interpolate between points separated by less than this distance, thus 200 local areas over the earth would probably suffice.

7 Conclusion

Five ensemble perturbation schemes have been tested on a simple model in an environment without model error. Although these tests are on a simple model, it is hoped that they will provide guidance on which schemes are likely to provide the best ensemble perturbation strategy for short range weather forecasting. These tests indicate that the ensemble Kalman filter is the best framework for generating initial condition perturbations for an ensemble. The localisation of the background error covariance is central to the performance of the EnKF. Combining either the ETKF or error breeding with random perturbations brings considerable improvements on any of these perturbation schemes alone.

The imbalance introduced by covariance localisation or adding noise does not appear to affect the results for this simple model. Although imbalance is an important issue in NWP models it has been shown [17] that covariance localisation may

be used without introducing serious imbalance. Model error must also be accounted for in NWP models, and one way of doing this is a "stochastic physics" parameterisation. Such schemes may improve the performance of the filtering perturbation methods (ETKF, error breeding) in the same way that combining these schemes with random perturbations has produced improved performance here. However, other authors [17] have found it difficult to devise a good model error parameterisation.

This study has been limited to a simple model without model error and with a size of the model state much less than that used for NWP models. However, the system is sufficiently large that it is not a trivial matter to provide an accurate ensemble forecast. This limitation on the system size may make the ETKF and EnKF appear more attractive than they would otherwise, since the size of the ensemble affects the determination of the perturbations. Model error (and its parameterisation) is likely to make the perturbation schemes appear more equal (since the ensemble quality will depend sensitively on the parameterisation quality).

Therefore, it is worth investigating the performance of the ETKF on atmospheric NWP models as this should prove to be a quick and effective method for generating ensemble perturbations. An EnKF which permits the localisation of background error covariances is likely to be superior to the simple ETKF, and a cheap method for achieving this has been shown to exhibit good performance.

Acknowledgements

The author wishes to thank Andy Lawrence for providing the initial code for the ETKF from which this study has its origin. Thanks also to Craig Bishop, Xugang Wang, Ken Mylne, Jim Hansen, Martin Leutbecher, Istvan Szunyogh and Tom Hamill for their advice and comments on this paper.

References

[1] C. H. Bishop, B. J. Etherton, and S. J. Majumdar. Adaptive sampling with the ensemble transform kalman filter. part 1: theoretical aspects. *Monthly Weather Review*, 129:420–436, 2001.

[2] S. M. Bozic. *Digital and Kalman Filtering*. Edward Arnold (Publishers) Ltd, Bedford Square, London, 1979.

[3] R. Buizza, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei, and Y. Zhu. Assessment of the status of global ensemble prediction. Submitted to Monthly Weather Review, 2004.

[4] R. Buizza, J Tribbia, F. Molteni, and T. Palmer. Computation of optimal unstable structures for a numerical weather prediction model. *Tellus A*, 45 No. 5:388–407, 1993.

[5] J. Du, S. L. Mullen, and F. Sanders. Short-range ensemble forecasting of quantitative precipitation. *Monthly Weather Review*, 125:2427–2459, 1997.

[6] G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte-carlo methods to forecast error statistics. *Journal of Geophysical Research - Oceans*, 99 C5:10143–10162, 1994.

[7] G. Evensen. Sampling strategies and square root analysis schemes for the enkf. http://www.nersc.no/ geir/EnKF/Publications/eve04a.pdf, 2004.

[8] G. Gaspari and S. E. Cohn. Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125 No. 554:723–757, 1999.

[9] T. M. Hamill, C. Snyder, and R. E. Morss. A comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles. *Monthly Weather Review*, 128:1835–1851, 2000.

[10] T. M. Hamill and C. Synder. A hybrid ensemble kalman filter-3d variational analysis scheme. *Monthly Weather Review*, 128 No. 8:2905–2919, 2000.

[11] P. L. Houtekamer and J. Derome. Methods for ensemble prediction. *Monthly Weather Review*, 123:2181–2196, 1995.

[12] P. L. Houtekamer, L. Lefaivre, and J. Derome. The rpn ensemble prediction system. In *Proceedings of the seminar on predictability*, volume II, pages 121–146, Reading, Berkshire, UK, 1995. ECMWF.

[13] P. L. Houtekamer and H. L. Mitchell. Data assimilation using an ensemble kalman filter technique. *Monthly Weather Review*, 126:796–811, 1998.

[14] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the AMSE - Journal of Basic Engineering*, 82 D:35–45, 1960.

[15] E. N. Lorenz. Predictability: a problem partly solved. In *Proceedings of the seminar on predictability*, volume I, pages 1–18, Reading, Berkshire, UK, 1995. ECMWF.

[16] E. N. Lorenz and K. A. Emanuel. Optimal sites for supplementary weather observations: simulation with a small model. *Journal of the Atmospheric Sciences*, 55:399–414, 1998.

[17] H. L. Mitchell, P. L. Houtekamer, and G. Pellerin. Ensemble size, balance, and model-error representation in an ensemble kalman filter. *Monthly Weather Review*, 130 No. 11:2791–2808, 2002.

[18] F. Molteni, R. Buizza, T. N. Palmer, and T. Petroliagis. The ecmwf ensemble prediction system: methodology and validation. *Quarterly Journal of the Royal Meteorological Society*, 122:73–119, 1996.

[19] E. Ott, B. R. Hunt, I. Szunyogh, A. V. Zimin, E. Kostelich, M. Corazza, E. Kalnay, D. J. Patil, and J. A. Yorke. A local ensemble kalman filter for atmospheric data assimilation. http://arxiv.org/abs/physics/0203058, in print Tellus, 2003.

[20] I. Szunyogh, E.J. Kostelich, G. Gyarmati, D.J. Patil, B.R. Hunt, E. Kalnay, E. Ott, and J.A. Yorke. Assessing a local ensemble kalman filter: perfect model experiments with the ncep global model. http://keck2.umd.edu/weather/weather_publications.htm, 2004.

[21] Z. Toth and E. Kalnay. Ensemble forecasting at nmc: the generation of perturbations. *Bulletin of the American Meteorological Society*, 74:2317–2330, 1993.

[22] Z. Toth and E. Kalnay. Ensemble forecasting at ncep and the breeding method. *Monthly Weather Review*, 125:3297–3319, 1997.

[23] X. Wang and C. H. Bishop. A comparison of breeding and ensemble transform kalman filter ensemble forecast schemes. *Journal of the Atmospheric Sciences*, 60:1140–1158, 2003.

[24] X. Wang, C. H. Bishop, and S. J. Julier. Which is better, and ensemble of positive/negative pairs or a centered spherical simplex ensemble? Submitted to Monthly Weather Review, 2003.

[25] M. Wei and Z. Toth. A new measure of ensemble performance: perturbation versus error correlation analysis (peca). *Monthly Weather Review*, 131:1549–1565, 2003.

[26] J. S. Whitaker and T. M. Hamill. Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, 130:1913–1924, 2002.